

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

F

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification: <b>G06F 12/00</b>		<b>A1</b>	(11) International Publication Number: <b>WO 00/07105</b>
			(43) International Publication Date: 10 February 2000 (10.02.2000)
(21) International Application Number: PCT/JP98/03335		<b>Published</b>	
(22) International Filing Date: 27 July 1998 (27.07.1998)			
(60) Parent Application or Grant HITACHI, LTD. [/]; O. YAMAGAMI, Kenji [/]; O. YAMAMOTO, Akira [/]; O. YAMAMOTO, Yasutomo [/]; O. WATANABE, Masaya [/]; O. YAMAGAMI, Kenji [/]; O. YAMAMOTO, Akira [/]; O. YAMAMOTO, Yasutomo [/]; O. WATANABE, Masaya [/]; O. OGAWA, Katsuo ; O.			
(54) Title: <b>COMPUTER SYSTEM</b> (54) Titre: <b>SYSTEME INFORMATIQUE</b>			
(57) Abstract <p>Data written in variable length format by a first computer is written in fixed length format into a physical volume by a controller according to a specified logic. A second computer reads out the data in fixed length format from the physical volume and extracts variable length data according to the same specified logic. The data written by the first computer can thereby be used by the second computer. Data read out by the second computer can be used similarly by the first computer.</p>			
(57) Abrégé <p>Les données écrites dans un format à longueur variable par un premier ordinateur sont écrites dans un format à longueur fixe par une unité de commande à l'intérieur d'un volume physique, en fonction d'un logique spécifique. Un second ordinateur lit les données d'un format à longueur fixe depuis ce volume physique, puis extrait les données d'une longueur variable en fonction de la même logique spécifique. Les données écrites par le premier ordinateur peuvent donc être utilisées par le second ordinateur. De même, les données lues par ce second ordinateur peuvent être utilisées par le premier ordinateur.</p>			

BEST AVAILABLE COPY

PCT

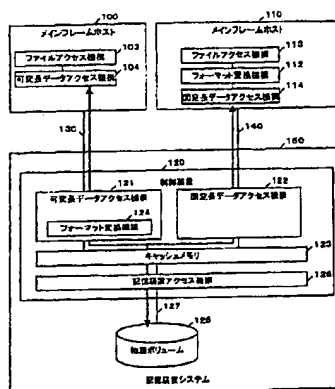
世界知的所有権機関  
国際事務局  
特許協力条約に基づいて公開された国際出願



(51) 国際特許分類6 G06F 12/00	A1	(11) 国際公開番号 WO00/07105
		(43) 国際公開日 2000年2月10日(10.02.00)
(21) 国際出願番号 PCT/JP98/03335	(74) 代理人 弁理士 小川勝男(OGAWA, Katsuo) 〒100-8220 東京都千代田区丸の内一丁目5番1号 株式会社 日立製作所内 Tokyo, (JP)	
(22) 国際出願日 1998年7月27日(27.07.98)		
<p>(71) 出願人 (米国を除くすべての指定国について) 株式会社 日立製作所(HITACHI, LTD.)(JP/JP) 〒101-8010 東京都千代田区神田駿河台四丁目6番地 Tokyo, (JP)</p> <p>(72) 発明者; および (75) 発明者/出願人 (米国についてのみ) 山神慈司(YAMAGAMI, Kenji)(JP/JP) 山本 彰(YAMAMOTO, Akira)(JP/JP) 山本康友(YAMAMOTO, Yasutomo)(JP/JP) 〒215-0013 神奈川県川崎市麻生区王禅寺1099番地 株式会社 日立製作所 システム開発研究所内 Kanagawa, (JP) 渡部真也(WATANABE, Masaya)(JP/JP) 〒259-1392 神奈川県秦野市堀山下1番地 株式会社 日立製作所 汎用コンピュータ事業部内 Kanagawa, (JP)</p>		<p>(81) 指定国 CN, JP, KR, US, 欧州特許 (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE)</p> <p>添付公開書類 国際調査報告書</p>

(54)Title: COMPUTER SYSTEM

(54)発明の名称 計算機システム



- |   |   |
|---|---|
| 100 ... Main frame host                       | 121 ... Variable length data access mechanism |
| 103 ... File access mechanism                 | 122 ... Fixed length data access mechanism    |
| 104 ... Variable length data access mechanism | 124 ... Format conversion mechanism           |
| 110 ... Main frame host                       | 123 ... Cache memory                          |
| 113 ... File access mechanism                 | 125 ... Storage device access mechanism       |
| 114 ... Fixed length data access mechanism    | 126 ... Physical volume                       |
| 120 ... Controller                            | 127 ... Storage system                        |

(57) Abstract

Data written in variable length format by a first computer is written in fixed length format into a physical volume by a controller according to a specified logic. A second computer reads out the data in fixed length format from the physical volume and extracts variable length data according to the same specified logic. The data written by the first computer can thereby be used by the second computer. Data read out by the second computer can be used similarly by the first computer.

(57)要約

第1の計算機が可変長フォーマットで書き出したデータを制御装置は所定の論理で固定長フォーマットで物理ボリュームに書き出す。また、第2の計算機で物理ボリュームからこの固定長フォーマットのデータを読み出し、同一の所定の論理で可変長データを抽出する。これにより、第1の計算機が書き出したデータを、第2の計算機において使用可能となる。さらに、同様に第2の計算機が書き出したデータを、第1の計算機において使用可能となる。

PCTに基づいて公開される国際出願のパブリック第一頁に掲載されたPCT加盟国を同定するために使用されるコード(参考情報)

AE アラブ首長国連邦	DM ドミニカ	KZ カザフスタン	RU ロシア
AL アルバニア	EE エストニア	LC セントルシア	SD スーダン
AM アルメニア	ES スペイン	LI リヒテンシュタイン	SE スウェーデン
AT オーストリア	FI フィンランド	LK スリ・ランカ	SG シンガポール
AU オーストラリア	FR フランス	LR リベリア	SI スロヴェニア
AZ アゼルバイジャン	GA ガボン	LS レソト	SK スロヴァキア
BA ボスニア・ヘルツェゴビナ	GB 英国	LT リトアニア	SL シェラ・レオネ
BB バルバドス	GD グレナダ	LU ルクセンブルグ	SN セネガル
BE ベルギー	GE グルジア	LV ラトヴィア	SZ スワジランド
BF ブルキナ・ファソ	GH ガーナ	MA モロッコ	TD チャード
BG ブルガリア	GM ガンビア	MC モナコ	TG トーゴ
BJ ベナン	GN ギニア	MD モルドヴァ	TJ タジキスタン
BR ブラジル	GW ギニア・ビサウ	MG マダガスカル	TZ タンザニア
BY ベラルーシ	GR ギリシャ	MK マケドニア旧ユーゴスラヴィア	TM トルクメニスタン
CA カナダ	HR クロアチア	共和国	TR トルコ
CF 中央アフリカ	HU ハンガリー	ML マリ	TT トリニダード・トバゴ
CG コンゴ	ID インドネシア	MN モンゴル	UA ウクライナ
CH スイス	IE アイルランド	MR モーリタニア	UG ウガンダ
CI コートジボアール	IL イスラエル	MW マラウイ	US 米国
CM カメルーン	IN インド	MX メキシコ	UZ ウズベキスタン
CN 中国	IS アイスランド	NE ニジェール	VN ヴェトナム
CR コスタ・リカ	IT イタリア	NL オランダ	YU ユーゴスラビア
CU キューバ	JP 日本	NO ノールウェー	ZA 南アフリカ共和国
CY キプロス	KE ケニア	NZ ニュー・ジーランド	ZW ジンバブエ
CZ チェコ	KG キルギスタン	PL ポーランド	
DE ドイツ	KP 北朝鮮	PT ポルトガル	
DK デンマーク	KR 韓国	RO ルーマニア	

**Description**

5

10

15

20

25

30

35

40

45

50

55

5

## 明 細 書

10

## 計算機システム

## 5 技術分野

15

本発明は、計算機システムに係り、特に、複数の計算機とこれらの計算機により共有される記憶装置とを含む計算機システムに関する。

## 背景技術

20

10 ホスト計算機の一部（特に大型計算機）は、可変長のデータフォーマットを持つ磁気ディスク装置を有する。これらのホスト計算機は、磁気ディスク装置にアクセスするために、CKDインタフェースと呼ばれる可変長データのアクセスインタフェースを有する。従来は可変長のレコードフォーマットがそのまま物理記憶装置上で実現されてきたが、近年

25 では、安価で比較的信頼性の低い複数の記憶装置にデータを分散配置し、データ転送の並列性を向上することによって性能を向上し、かつデータに冗長度を持たせることによって高い信頼性を得るRAIDと呼ばれる手法が一般的になりつつある。この手法は米国、バークレイ、カリフォルニア大学のパターンソン（Patterson）等が、米国計算機学会データ管理専門グループ技術報告（Proc. ACM SIGMOD）1988年6月号において、「安価なディスクの冗長アレイ（RAID）のケース（A Case for Redundant Array of Inexpensive Disks）」を発表して以来、

30 急速に広まりつつある。

35

25 文献：「メインフレーム '98」，日経 BP 社，p126-130に記述されるように、RAID技術を適用した記憶装置システムでは、FBA (Fixed Block Architecture) フォーマットと呼ばれる固定長のデータフォーマットを持つ磁気ディスク装置が用いられていて、このディスク装置にアクセスするインタフェースとして、SCSIインタフェース（あるいはSCSIプロトコル）が用いられる。したがって、ホスト計算機からはCKDインタフェースによって可変長データへアクセスする

40

45

50

55

。そして、内部にFBA形式の磁気ディスクを用いた記憶装置システムでは、データの書き込みの場合には可変長フォーマットから固定長フォーマットへ、また、データの読み出しの場合には固定長フォーマットから可変長フォーマットへ変換するフォーマット変換機構を備えている。

5      このようなフォーマット変換機構の一例として、例えば特開平6-150557が知られている。

15      近年の計算センタでは、ホスト計算機とUNIXサーバ、PCサーバ等のオープンホストとが混在してシステム構築されることが多い。このような構成においては、記憶装置システムの管理を容易にしてトータルストレージコストを低減する要求が高い。これに応えるために、例えば「メインフレーム'98」、日経BP社、p145あるいはp151に見られるような、可変長のアクセスインタフェースと固定長のアクセスインタフェースの双方を備える記憶装置システムが開発されている。

25      一方、最近では、ホスト計算機の内蔵ディスク装置に、SCSIインタフェースを持つものが出現してきている。例えば、「メインフレーム'98」、日経BP社、p53-54に記載されているサーバでは、FBAフォーマットのディスク装置をホスト計算機の筐体に内蔵し、OSが生成した可変長データアクセスのコマンド群(CCW: Channel Command Word チェイン)をSAPと呼ばれるプロセッサが解釈して、  
30      固定長データアクセスのコマンド群(SCSIコマンド)に変換して、内蔵ディスクへの入出力処理を実行する。この内蔵ディスクでは可変長のデータを固定長のフォーマットへ埋め込んでおり、SAPでフォーマット変換を行っている。このようにOSが生成したCCWをSAPで  
35      SCSIコマンドに変換するメリットは、OSあるいはアプリケーションプログラムの変更が不要な点である。

45      前記サーバでは、内蔵ディスクに関するものだが、本従来技術を適用すれば、ホスト計算機に固定長のアクセスインタフェースを有する外部記憶装置システムを接続することも可能になる。このような外部記憶装置は多数のベンダから提供されているので、ユーザは価格、性能、信頼  
50      性に依じて幅広い選択が可能になり、その利点は非常に大きい。

55      外部記憶装置は複数のホスト計算機から共有されることが前提であり

、かつそのことが記憶装置をホスト計算機の筐体外に持つことの意義でもある。しかしながら、上記従来技術はもともと内蔵ディスクに関するものなので、データを共有することを考慮していない。

ここで、データを共有する場合の課題を以下に説明する。

- 5 「メインフレーム '98」、日経BP社、p 145あるいはp 151  
15 に記載されているような固定長のアクセスインタフェースと可変長のアクセスインタフェースの両方を備える記憶装置システムに、固定長のアクセスインタフェースを有するホスト計算機と、可変長のアクセスインタフェースを有するホスト計算機とを接続しても、固定長のアクセス  
20 10 インタフェースを有するホスト計算機と記憶装置システム間で、データのフォーマット、つまり可変長フォーマットと固定長フォーマットの変換方法、が共通でなければ、データを共有できない。つまり、可変長アクセスインタフェースを持つホスト計算機から記憶装置にライトされたデータを、固定長アクセスインタフェースを持つメインフレームから読出  
25 15 す、あるいはその逆はできないという課題がある。

本発明の目的は、固定長のアクセスインタフェースを有するホスト計算機と可変長のアクセスインタフェースを有するホスト計算機間で、外部記憶装置に格納されたデータを共有可能とすることである。

#### 20 発明の開示

- 35 上記目的を達成するために、本発明の計算機システムは、可変長フォーマットによりフォーマットされた第1の論理ボリュームへのアクセスインタフェースを有する第1の計算機と、固定長フォーマットによりフ  
40 ォーマットされた第2の論理ボリュームへのアクセスインタフェースと  
25 、前記第2の論理ボリュームから読み出した固定長データから所定の論理に従って可変長データを抽出するとともに、可変長データを前記所定の論理に従って前記第2の論理ボリュームに書き込む固定長データに変換する手段とを有する第2の計算機と、固定長フォーマットによりフ  
45 ォーマットされた物理ボリュームと、前記第1及び第2の計算機に接続さ  
30 れ、前記第1の論理ボリュームを前記所定の論理に従って固定長フォーマットによりフォーマットされた前記物理ボリュームにフォーマット変  
50

換し、前記第2の論理ボリュームと前記物理ボリュームのアドレスを一致させて対応付ける制御装置とを有する。

第1の計算機が可変長フォーマットで書き出したデータを制御装置は所定の論理で固定長フォーマットで物理ボリュームに書き出す。また、第2の計算機で物理ボリュームからこの固定長フォーマットのデータを読み出し、前記所定の論理で可変長データを抽出する。これにより、第1の計算機が書き出したデータを、第2の計算機において使用可能となる。

また、第2の計算機が所定の論理に従って可変長データを固定長データに変換して書き出したデータを、制御装置は固定長フォーマットで物理ボリュームに書き出す。また、制御装置は、第1の計算機の要求に従って物理ボリュームからこの固定長フォーマットのデータを読み出し、前記所定の論理で可変長データに変換して第1の計算機にこの可変長データを渡す。これにより、第2の計算機が書き出したデータを、第1の計算機において使用可能となる。

#### 図面の簡単な説明

第1図は、第一の実施例における計算機システムの構成図である。

第2図は、第一の実施例の概要説明図である。

第3図は、記憶装置の構成図である。

第4図は、可変長データのトラックフォーマットである。

第5図は、固定長データのトラックフォーマットである。

第6図は、可変長データと固定長データのフォーマット変換方法の説明図である。

第7図は、制御装置およびホスト計算機が有するフォーマット変換機構のフローチャートである。

第8図は、ホスト計算機が有するフォーマット変換機構におけるライト処理のフローチャートである。

第9図は、制御装置が有する固定長データアクセス機構のフローチャートである。



発明を実施するため最良の形態

以下、本発明の第一の実施例を図を用いて説明する。

第1図は本発明における計算機システムの構成図を示す。本発明では、可変長のアクセスインタフェースを有する一台以上のホスト計算機100と、固定長のアクセスインタフェースを有する一台以上のホスト計算機110と、一台以上の記憶装置システム150から構成される。

ホスト計算機100には、アプリケーションプログラムからの入出力要求を受けて、可変長フォーマットの記憶装置にアクセスするためのコマンド群(CCWチェーン)を生成するファイルアクセス機構103と、ファイルアクセス機構103が生成したCCWチェーンを実行し、記憶装置システム150との間で入出力処理を行う可変長データアクセス機構104が存在する。

一方、ホスト計算機110は、アプリケーションプログラムからの入出力要求を受けて、可変長フォーマットの記憶装置にアクセスするためのCCWチェーンを生成するファイルアクセス機構113と、CCWチェーンを解釈して、可変長データと固定長データ間のフォーマット変換を行い、固定長の記憶装置へアクセスするためのSCSIコマンドを生成するフォーマット変換機構112と、フォーマット変換機構112が生成したSCSIコマンドを実行し、入出力処理を行う固定長データアクセス機構114を有する。

記憶装置システム150は、データを格納する物理ボリューム128と、物理ボリューム128とホスト計算機100または110間で入出力処理を制御する制御装置120から構成される。

物理ボリューム128は固定長フォーマットの記憶装置であり、記憶装置バス127により制御装置120に接続されている。制御装置120は記憶装置アクセス機構126を用いて物理ボリューム128への入出力処理を実行する。物理ボリューム128の実例は、SCSIインタフェース、あるいはファイバチャネルインタフェースを有するFBAディスクである。

制御装置120とホスト計算機100はホストバス130により接続されていて、ホスト計算機100から発行されたCCWチェーンは、可

変長データアクセス機構 121 により解釈され、可変長ボリュームの入出力処理が実行される。本発明では、ホストバス 130 の実例は ESCON あるいはファイバチャネルであり、可変長ボリュームは CKD ディスクを想定している。

一方、制御装置 120 とホスト計算機 110 とはホストバス 140 により接続されていて、ホスト計算機 110 から発行された SCSI コマンドは、固定長データアクセス機構 122 により解釈され、固定長ボリュームへの入出力処理が実行される。本発明では、ホストバス 130 の実例は SCSI バスやファイバチャネルを想定していて、固定長ボリュームは FBA ディスクを想定している。

キャッシュメモリ 123 は、物理ボリューム 128 上のデータを一時的に格納するためのメモリである。

物理ボリューム 128 は固定長フォーマットなので、可変長データアクセス機構 121 では、固定長フォーマットと可変長フォーマットの間でフォーマット変換を実行する必要がある。この処理はフォーマット変換機構 124 にて実行される。

ホスト計算機 100 はアクセス対象の可変長ボリュームを論理ボリューム 200 a として認識しており、ホスト計算機 110 は、アクセス対象の固定長ボリュームを論理ボリューム 200 b として認識している。

第 2 図を用いてこれらの様子を説明する。

ホスト計算機 100 が認識する論理ボリューム 200 a は、可変長フォーマットによりフォーマットされた記憶装置であり、目的のデータにアクセスする場合には、ホスト計算機 100 は、可変長データアクセス機構 104 を用いて、論理ボリューム 200 a の番号と、アクセスするデータの論理ボリューム 200 a 上のアドレス等を、可変長ボリュームのアクセスインタフェースにしたがって指定する。制御装置 120 は、フォーマット変換機構 124 によって、論理ボリューム 200 a の可変長データを固定長のデータフォーマットを有する物理ボリューム 128 にフォーマット変換する。すなわち、ホスト計算機 100 からアクセス要求があると、指定された論理ボリューム 200 a の番号と、アクセスアドレスを、対応する物理ボリューム 128 上のアドレスに変換する。

5  
10  
なお、フォーマット変換機構 124 によって実現される、可変長データと固定長データ間のフォーマット変換方法とアドレスの対応付けについては後述する。

5  
15  
20  
25  
30  
一方、ホスト計算機 110 が認識する論理ボリューム 200b は、固定長フォーマットでフォーマットされた記憶装置であり、目的のデータにアクセスする場合には、ホスト計算機 110 は、固定長データアクセス機構 114 を用いて論理ボリューム 200b の番号と、論理ボリューム 200b 上のアドレス等を、固定長ボリュームのアクセスインタフェースにしたがって指定する。実はこの論理ボリューム 200b は物理ボリューム 128 そのものである。つまり、論理ボリューム 200b 上のアドレスと物理ボリューム 128 上のアドレスは一致している。ホスト計算機 110 がデータを読み出す場合は、固定長データアクセス機構 114 は論理ボリューム 200b、すなわち物理ボリューム 128 から固定長データを読み出し、フォーマット変換機構 112 により、可変長データを抽出する。またデータを書込む場合には、フォーマット変換機構 112 により可変長データを固定長データへ変換し、固定長データアクセス機構 114 により、固定長データを論理ボリューム 200b、すなわち物理ボリューム 128 へ書込む。

35  
40  
45  
ここでのポイントは、フォーマット変換機構 124 と 112 が、同一方法によってフォーマット変換を行うことである。こうすることで、一方のフォーマット変換機構によって変換された固定長データを、他方のフォーマット変換機構により可変長に復元することができる。

50  
55  
また、本発明で扱う記憶装置システム 150 に RAID 技術を適用して、複数の記憶装置の一部領域から、一台の物理ボリューム 128 を構成することもできる。この場合には物理ボリューム 128 上のアドレスを RAID を構成する記憶装置にマッピングする必要があるが、これには、例えば特開平 5-197498 に記載された方法を適用すればよい。

60  
65  
なお、論理ボリューム 200a および 200b、物理ボリューム 128 には一意に番号が付けられていて、論理ボリューム 200a あるいは論理ボリューム 200b の番号が指定されると、対応する物理ボリューム

1 2 8 が一つ定まる。

次に、第 3 図を用いて論理ボリューム 2 0 0、物理ボリューム 1 2 8  
の一構成例について、説明する。各ボリュームとも基本的な構成は同じ  
であり、次の説明では、これらのボリュームを総称して記憶装置と呼ぶ  
5 。

記憶装置 4 0 0 は回転する記憶媒体と、複数のディスク 4 1 0 と、デ  
ィスク 4 1 0 の枚数に等しい数のヘッド 4 1 3 から構成される。ディス  
ク 4 1 0 が一回転した時に、一つのヘッド 4 1 3 がアクセスできるディス  
ク 4 1 0 の領域をトラック 4 1 1 と呼び、全てのヘッド 4 1 3 の下を  
通過するトラック 4 1 1 の円筒状の集合をシリンダ 4 1 2 と呼ぶ。シリ  
ンダ 4 1 2 は最外周から内周に向かって 0 から昇順に番号が付けられて  
いて、ヘッド 4 1 3 にも上から下へ 0 から昇順に番号が付けられている  
。したがって、シリンダ番号とヘッド番号の組みで、トラック 4 1 1 を  
特定できる。トラック 4 1 1 上にはインデックス 5 0 4 と呼ばれる特別  
なマークが付与されていて、これによりトラック 2 1 1 の先頭を認識す  
る。

論理ボリューム 2 0 0 と、物理ボリューム 1 2 8 は、上述した基本構  
成は同じだが、可変長フォーマットか固定長フォーマットかによって、  
記憶装置 4 0 0 上のアドレス表現およびデータの格納フォーマットが異  
なる。まず第 4 図を用いて可変長ボリュームである論理ボリューム 2 0  
0 a の場合のデータ格納フォーマットを説明する。

トラック 4 1 1 は一個のホームアドレス（以下、H A と略す）5 1 0  
と一個以上のレコード 5 0 0 から構成される。H A 5 1 0 は各トラック  
4 1 1 の先頭に位置する最初のフィールドである。各レコード 5 0 0 は  
一個以上のフィールド、すなわちカウント部 5 0 1、データ部 5 0 2、  
さらに場合によっては、データ部の前にキー部（図示せず）が存在しう  
る。カウント部 5 0 1 は固定長のフィールドで、そのレコード 5 0 0 の  
アドレス、後続するフィールド（上記のデータ部 5 0 2 とキー部）の長  
さなどを格納している。各レコード 5 0 0 は可変長であること、つまり  
データ部 5 0 2 の長さがレコード 5 0 0 毎に異なっても良いこと、  
が許されている。このために、カウント部 5 0 1 にそのレコードのキー

部およびデータ部 502 の長さが格納されていて、カウント部 501 を  
見れば、そのレコード 500 の各フィールドの長さが分かるようになって  
いる。レコード 500 のアドレスは、シリンダ番号、ヘッド番号、レ  
コード番号、すなわちトラック先頭から付与されたシーケンシャルな番  
5 号の組みで示す。

次に第 5 図を用いて固定長ボリュームである論理ボリューム 200b  
、物理ボリューム 128 のアドレス表現について説明する。固定長の記  
憶装置 400 では、各トラック 411 はあらかじめ定められた固定長の  
領域（以降、ブロック 600 と呼ぶ）から構成されていて、各ブロック  
20 600 は記憶装置 400 内で一意に番号付けされている。従って記憶装  
置 128 のある領域にアクセスする際には、その領域の先頭のブロック  
600 の番号と、それに後続するブロック 600 の個数を指定すれば良  
い。

続いて、可変長の論理ボリューム 200a を固定長の物理ボリューム  
15 にフォーマット変換する方法を第 6 図を用いて説明する。

フォーマット変換に当って、可変長フォーマットで表現された各トラ  
ック 411 を固定長のブロック 600 の大きさと等しい領域に分割し、  
番号 0 で示される先頭のブロック 600 に、シリンダ番号 0、ヘッド番  
号 0 で示される先頭トラック 411 の先頭領域を割り当てる。以降、後  
20 続するブロック 600 にトラック 411 上の後続する領域を割り当てて  
いき、先頭のトラック 411 の全領域が割り当てられると、後続するブ  
ロック 600 には次トラック、すなわちシリンダ番号 0、ヘッド番号 1  
のトラック 411 の先頭領域を割り当てる。以下同様にして、各トラッ  
ク 411 の全領域を各ブロック 600 に対応付ける。

25 各トラックの長さは論理ボリューム 200a のタイプによって固定的  
に定められているので、以下の数式を用いれば、トラックアドレスから  
そのトラック 411 を格納している先頭のブロック 600 の番号と、ト  
ラック当りのブロック 600 の個数を求めることができる。

（数 1）                      トラック当たりのブロック数 = <トラック容量 ÷ ブロ  
30 ック容量>

（数 2）                      トラック先頭のブロック番号 = <シリンダ番号 × ヘッ

5

10

15

20

25

30

35

40

45

50

55

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55

ド数+ヘッド番号)×トラック当たりのサブブロック数

ここで $x$ は、 $x$ 以上の最小の整数である。

例えば、サブブロック340の容量が512バイト、トラック211の容量が47キロバイト、ヘッド数は15とすると、トラック当たりのサブブロック数は94であり、シリンダ番号0、ヘッド番号1のトラック411の先頭サブブロックは94、シリンダ番号100、ヘッド番号5のトラック411の先頭サブブロックは141、470となる。

上記の方法によって、論理ボリューム200aの全ての領域を、物理ボリューム128に割り当てるが、ここで示したフォーマット変換はあくまでも一例である。一般的にはさまざまなフォーマット変換方法が考えられるが、本発明で重要なことは、ある一つのフォーマット変換を、全てのホスト計算機110が有するフォーマット変換機構112、および制御装置120が有するフォーマット変換機構124が認識し、共通に使用することである。

次に第7図を用いて、ホスト計算機100からの入出力要求に対する制御装置120の処理を説明する。制御装置120がホスト計算機100から入出力要求を受領すると、指定された論理ボリューム200aに対応する物理ボリューム128の番号を求める(ステップ880)。次にステップ800で、(数1)および(数2)を使って、アクセス対象トラックが格納されている物理ボリューム128上のブロック番号とブロック数を求める。次にステップ810で、キャッシュ123に目的とするトラック411が存在するかどうかを調べ、もし存在していなければ、当該トラック411を物理ボリューム128からキャッシュ123へロードする(ステップ820)。

もしキャッシュ123上に目的とするトラック411が存在していたら、ステップ820はスキップされる。続いてステップ830で、当該トラック411の先頭レコード500のカウント部501へアクセスし、ステップ840でアクセス対象のレコード500かどうかを調べる。ここでステップ840は次の方法により実現する。すなわち制御装置120はホスト計算機100から指示されたアクセス対象のアドレス、つまりシリンダ番号、ヘッド番号、レコード番号を初期値として、次にア

5  
10  
アクセスすべきアドレスを保持していて、カウント部 501 から読出した  
アドレスと比較する。比較の結果もし一致していればアクセス対象のレ  
コード 500 であり、次アクセス対象のアドレスに設定する。

5  
15  
もし当該レコード 500 がアクセス対象なら、ステップ 850 でホス  
ト計算機 100 から指定されたフィールドへアクセスする。すなわちもし  
リードアクセスなら可変長データアクセス機構 121 を使用してキャ  
ッシュ 123 からホスト計算機 100 へ指定フィールドを転送し、ライ  
トアクセスならホスト計算機 100 からキャッシュ 123 へ指定フィー  
ルドへの書込みを行う。

20  
25  
16  
30  
指定フィールドへのアクセスが完了すると、ホスト計算機 100 が後  
続するレコード 500 へのアクセスを要求しているかどうかを、CCW  
チェーンにより判定する（ステップ 860）。もしそうなら、次レコー  
ド 500 のカウント部 501 を読出す。同様に、ステップ 840 で当該  
レコード 500 がアクセス対象でなければ、当該レコード 500 を読み  
飛ばし、ステップ 870 で次レコード 500 のカウント部 501 を読出  
す。その後、再びステップ 840 に戻り、当該レコード 500 がアクセ  
ス対象かどうかを調べる。

ステップ 840 からステップ 870 までの処理をホスト計算機 100  
からのアクセス要求が終了するまで繰り返す。

35  
40  
25  
ホスト計算機 100 からのアクセス要求がライトだった場合は、キャ  
ッシュ 123 上にライトデータが格納されている。このため、当該デー  
タをキャッシュ 123 から物理ボリューム 128 にライトしなくてはな  
らないが、この実現も容易である。すなわち、ステップ 800 で求めた  
ように、（数 1）および（数 2）を使用してトラック 411 のアドレス  
を物理ボリューム 128 上のアドレス、つまり先頭ブロック番号とブロ  
ック数を算出する。そのアドレスにしたがって記憶装置アクセス機構 1  
26 を使用して、物理ボリューム 128 へ書込めばよい。

45  
50  
55  
以上、ホスト計算機 100 から論理ボリューム 200 a へのアクセス  
方法について説明した。以降では、ホスト計算機 100 と 110 と間で  
記憶装置 400 を共有するための方法について説明する。このために、  
ホスト計算機 110 は物理ボリューム 128 を論理ボリューム 200 b

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55

として認識しているため、論理ボリューム 200b 上のアドレスと物理  
ボリューム 128 上のアドレスは一致している。ホスト計算機 110 は  
、この物理ボリューム 128 にアクセスして固定長のデータを読み出し、  
フォーマット変換機構 112 によって、読み出した固定長のデータから目  
5 的とする可変長のレコード 500 を抽出する。ホスト計算機 110 内の  
フォーマット変換機構 112 は、制御装置 120 のフォーマット変換機  
124 が実現する可変長フォーマットから固定長フォーマットへの埋  
め込み方法を知っているため、フォーマット変換機構 124 と同様の処  
理を実行することによって、固定長データから目的の可変長レコードへ  
10 アクセスすることができる。以下、第 7 図、第 8 図を用いてこの実現方  
法の詳細について説明する。

フォーマット変換機構 112 は、ファイルアクセス機構 113 で作成  
された CCW チェインを受け取ると、その CCW チェインを解釈してリ  
ードアクセスか、ライトアクセスかを判断する。まず、第 7 図を用いて  
15 リードアクセスだった場合の処理について説明する。論理ボリューム 2  
00b と物理ボリューム 128 の対応付けは制御装置 120 が実行する  
ので、論理ボリューム 200b に対するアクセス要求の場合、ステップ  
880 は不要になる。次にステップ 800 では CCW チェインから、ア  
クセス対象のアドレス、すなわちシリンダ番号、ヘッド番号、レコード  
20 番号を求めて、これらより（数 1）と（数 2）を使って、アクセス対象  
のトラック 411 が格納されているブロック番号とブロック数を求める  
。次にステップ 810 で、目的とするトラック 411 がホスト計算機 1  
10 上のキャッシュ（図示せず）上に格納されているかどうかを判定す  
る。ここでキャッシュとは、ホスト計算機 110 上に存在しているメモ  
25 リであり、例えばこの目的のためだけに定義されたバッファでも良いし  
、あるいは主記憶の一部を使用してこれを実現してもよい。もし目的の  
トラック 411 がキャッシュ上に存在しない場合は、ステップ 820 で  
記憶装置システム 150 にアクセスして、当該トラック 411 をキャッ  
シュへ読み出す。トラック 411 を記憶装置システム 150 から読み出す際  
30 には、ステップ 800 で求めたブロック番号とブロック数を基に読み出  
しの SCSI コマンドを生成し、固定長データアクセス機構 114 にこれ



5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55

を実行させる。以降、フォーマット変換機構 112 は、CCW チェインの内容に従い、ステップ 830 から 870 までを実行するが、これは前記の説明と同様なので省略する。

次にフォーマット変換機構 112 がファイルアクセス機構 113 が作成した CCW チェインを解析した結果、ライトアクセスだった場合の処理を、第 8 図を用いて説明する。第 8 図において、ステップ 900、930、940、950、960、970 の処理は、第 7 図におけるステップ 800、830、840、850、860、870 の処理と全く同じなので省略し、残りのステップについて説明する。

10  
15  
20  
25  
30  
35  
40  
45  
50  
55

ステップ 920 では、ステップ 900 で求めたブロック番号とブロック数により特定される領域を、固定長データアクセス機構 114 を介して、記憶装置システム 150 から読出す。次にステップ 930 からステップ 970 で、CCW チェインを解析して、要求された全てのレコード 300 の更新がキャッシュ上で完了すると、ステップ 980 で、固定長データアクセス機構 114 を介して再び記憶装置システム 150 へ当該トラック 411 を書込む。

次に、ホスト計算機 110 からアクセスがあった際の制御装置 120 の処理を第 9 図を用いて説明する。制御装置 120 の処理のポイントは、物理ボリューム 128 に格納されているデータを、フォーマット変換を行うことなく、そのまま SCSI コマンドによりアクセスさせることである。ホスト計算機 110 はフォーマット変換機構 112 によって、記憶装置システム 150 から読出した固定長データから可変長データを抽出し、また可変長データを固定長データへ変換して、記憶装置システム 150 へ書込む。

25  
30  
35  
40  
45  
50  
55

まずステップ 1000 で、ホスト計算機 110 から指定された論理ボリューム 200b に対応する物理ボリューム 128 を求める。そして、ホスト計算機 110 から指定された論理ボリューム 200b に対するアクセスアドレス、つまりブロック番号とブロック数を、物理ボリューム 128 のアクセスアドレスとする。次にステップ 1010 でリードアクセスかどうかを判定し、もしそうであればステップ 1020 で、目的のデータがキャッシュ 123 上に存在するかどうかを調べ、存在しなければ

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55

ばステップ1030で記憶装置アクセス機構126を使用して目的とするデータをキャッシュ123へロードする。その後、ステップ1040で、リードデータをホスト計算機110へ転送し、処理を完了する。もしステップ1010でライトアクセスと判定された場合は、ホスト計算機110からライトデータを受領し、キャッシュ123へ格納して処理を完了する。その後、キャッシュ123から物理ボリューム128へデータを書込む必要があるが、この際にはホスト計算機110から指定された書込みアドレスがそのまま物理ボリューム128上のアドレスであるため、論理ボリューム200bの番号を物理ボリューム128の番号に番号に変換したのち、そのアドレスへ書込めば良い。

以上の処理によれば、ホスト計算機100からライトされた可変長のデータは、フォーマット変換機構124により固定長データに変換されて、物理ボリューム128へ書込まれる。そのデータをホスト計算機110から読出す場合には、当該データが格納されている物理ボリューム128上のアドレスをホスト計算機110が指定し、固定長データのままで読出す。そしてフォーマット変換機構112によって、目的とする可変長のデータを抽出する。

一方、ホスト計算機110がデータを書込む際には、フォーマット変換機構112により、可変長データを固定長データへ変換して、物理ボリューム128へ書込む。そのデータをホスト計算機100が読出す際には、当該データが格納されている論理ボリューム200a上のアドレスを指定し、制御装置120は指定されたアドレスを物理ボリューム128上のアドレスに変換して、固定長データを読出す。そしてフォーマット変換機構124により、可変長データを抽出して、ホスト計算機へ当該データを転送する。

これらの処理によって、可変長のアクセスインタフェースを持つホスト計算機100と、固定長のアクセスインタフェースを持つホスト計算機110間で、記憶装置システム150に格納されたデータを共有することが可能になる。

産業上の利用可能性

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55

以上のように、本発明にかかる複数のホスト計算機におけるデータ共有方法は、アクセスインタフェースが異なる複数のホスト計算機によるデータの共有方法として有用であり、固定長のアクセスインタフェースを有するホスト計算機と可変長のアクセスインタフェースを有するホスト計算機との間で、固定長フォーマットでフォーマットされた記憶装置のデータを共有する計算機システムを構築するのに適している。

**Claims**

5

10

15

20

25

30

35

40

45

50

55

## 請 求 の 範 囲

1. 可変長フォーマットによりフォーマットされた第1の論理ボリュームへのアクセスインタフェースを有する第1の計算機と、

固定長フォーマットによりフォーマットされた第2の論理ボリュームへのアクセスインタフェースと、前記第2の論理ボリュームから読み出した固定長データから所定の論理に従って可変長データを抽出するとともに、可変長データを前記所定の論理に従って前記第2の論理ボリュームに書き込む固定長データに変換する手段とを有する第2の計算機と、

固定長フォーマットによりフォーマットされた物理ボリュームと、

前記第1、第2の計算機及び前記物理ボリュームに接続され、前記第1の論理ボリュームを前記所定の論理に従って固定長フォーマットによりフォーマットされた前記物理ボリュームにフォーマット変換し、前記第2の論理ボリュームと前記物理ボリュームのアドレスを一致させて対応付ける制御装置とから構成される計算機システム。

2. 可変長フォーマットによりフォーマットされた第1の論理ボリュームへの書き出しインタフェースを有する第1の計算機と、

固定長フォーマットによりフォーマットされた第2の論理ボリュームへの読み出しインタフェースと、前記第2の論理ボリュームから読み出した固定長データから所定の論理に従って可変長データを抽出する手段とを有する第2の計算機と、

固定長フォーマットによりフォーマットされた物理ボリュームと、

前記第1、第2の計算機及び前記物理ボリュームに接続され、前記第1の論理ボリュームを前記所定の論理に従って固定長フォーマットによりフォーマットされた前記物理ボリュームにフォーマット変換し、前記第2の論理ボリュームと前記物理ボリュームのアドレスを一致させて対応付ける制御装置とから構成される計算機システム。

3. 可変長フォーマットによりフォーマットされた第1の論理ボリュームへの読み出しインタフェースを有する第1の計算機と、

固定長フォーマットによりフォーマットされた第2の論理ボリュームへの書き出しインタフェースと、可変長データを前記所定の論理に従って前記第2の論理ボリュームに書き込む固定長データに変換する手段と

を有する第 2 の計算機と、

固定長フォーマットによりフォーマットされた物理ボリュームと、

前記第 1、第 2 の計算機及び前記物理ボリュームに接続され、前記第 1 の論理ボリュームを前記所定の論理に従って固定長フォーマットによりフォーマットされた前記物理ボリュームにフォーマット変換し、前記第 2 の論理ボリュームと前記物理ボリュームのアドレスを一致させて対応付ける制御装置とから構成される計算機システム。

4. 可変長フォーマットによりフォーマットされた第 1 の論理ボリュームへのアクセスインタフェースを有する第 1 の計算機と、

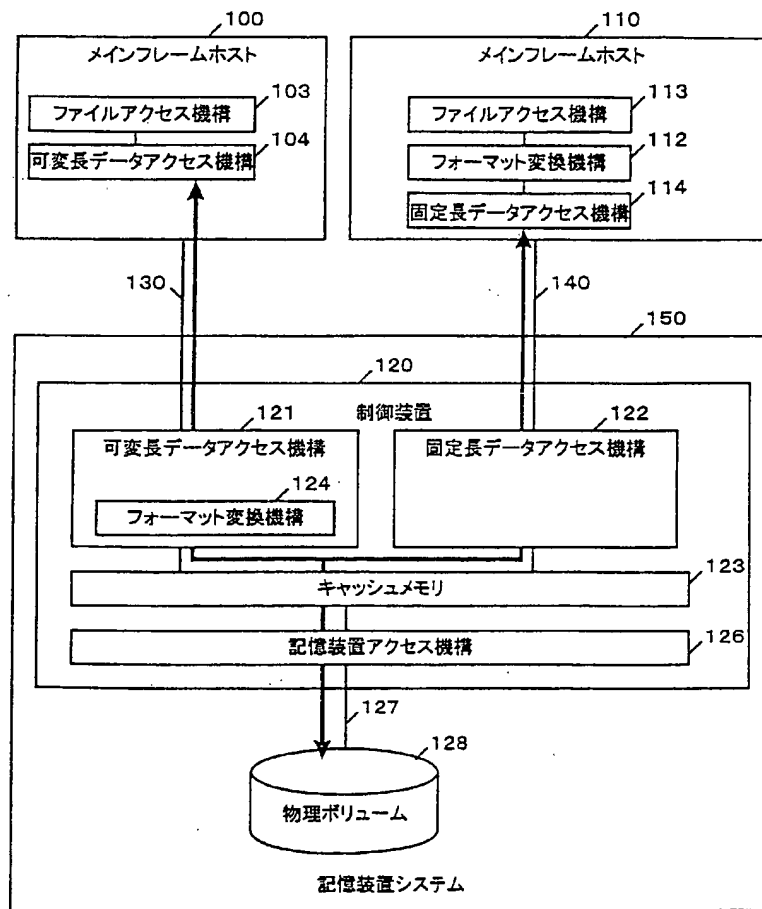
10 固定長フォーマットによりフォーマットされた第 2 の論理ボリュームへのアクセスインタフェースと、前記第 2 の論理ボリュームから読出した固定長データから可変長データを抽出するとともに、可変長データを固定長フォーマットに変換し前記第 2 の論理ボリュームに書込む可変長 - 固定長フォーマット変換機能を有する第 2 の計算機と、

15 固定長フォーマットによりフォーマットされた物理ボリュームと、

前記第 1、第 2 の計算機及び前記物理ボリュームに接続され、前記可変長 - 固定長フォーマット変換機能を有し、これを用いて前記第 1 の論理ボリュームのデータを前記物理ボリュームに格納するとともに、前記第 2 の論理ボリュームのアドレスと前記物理ボリュームのアドレスとを  
20 対応付けることによって、前記第 2 の計算機からの前記物理ボリュームへのアクセスを受け付ける制御装置とから構成される計算機システム。

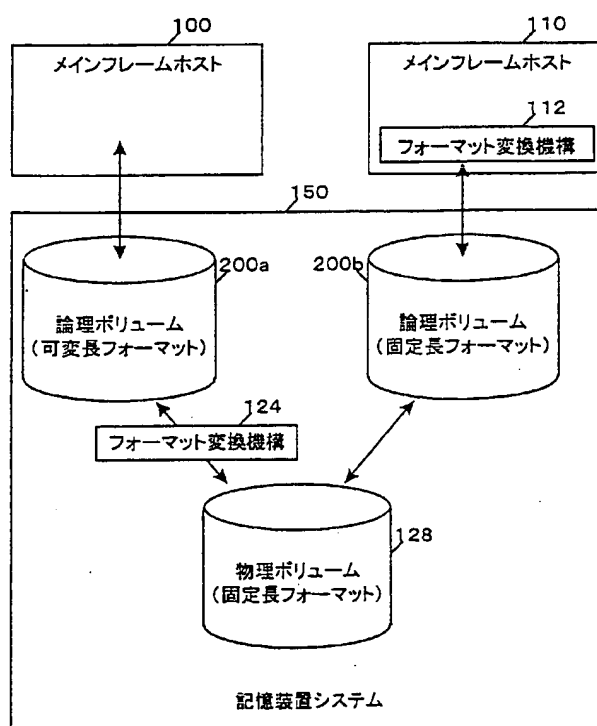
1 / 9

第 1 図



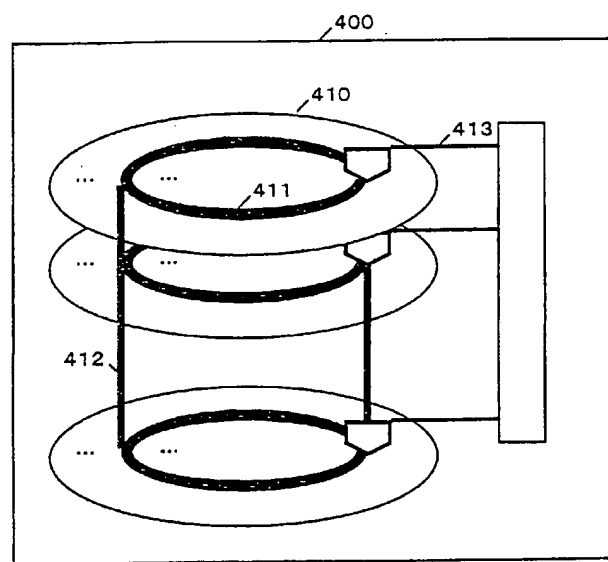
2 / 9

第 2 図

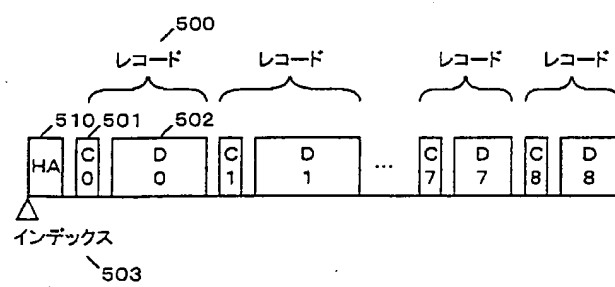




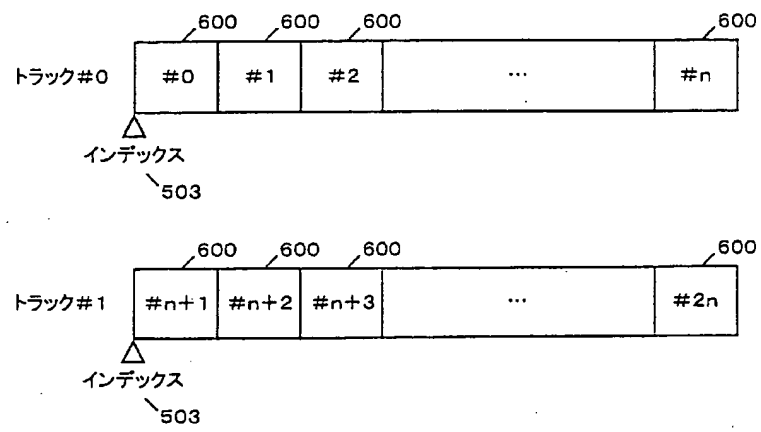
第 3 図



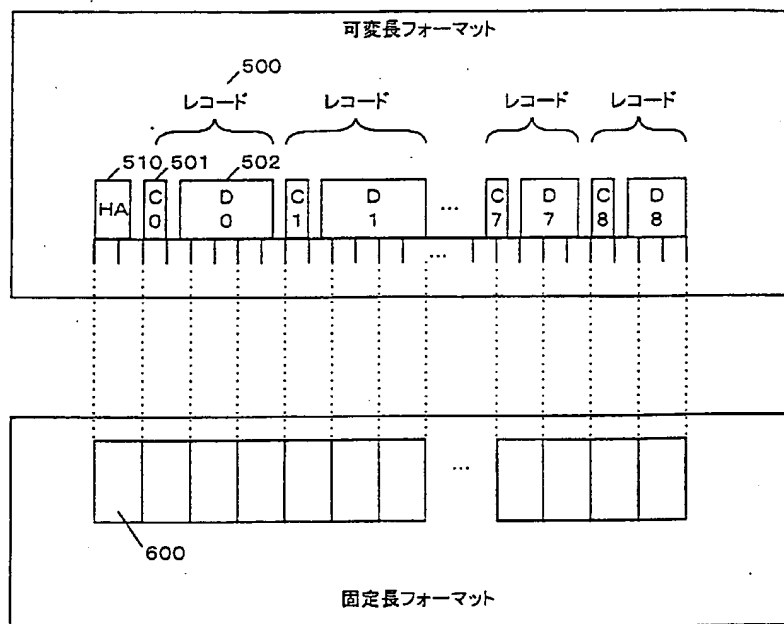
第 4 図



第 5 図

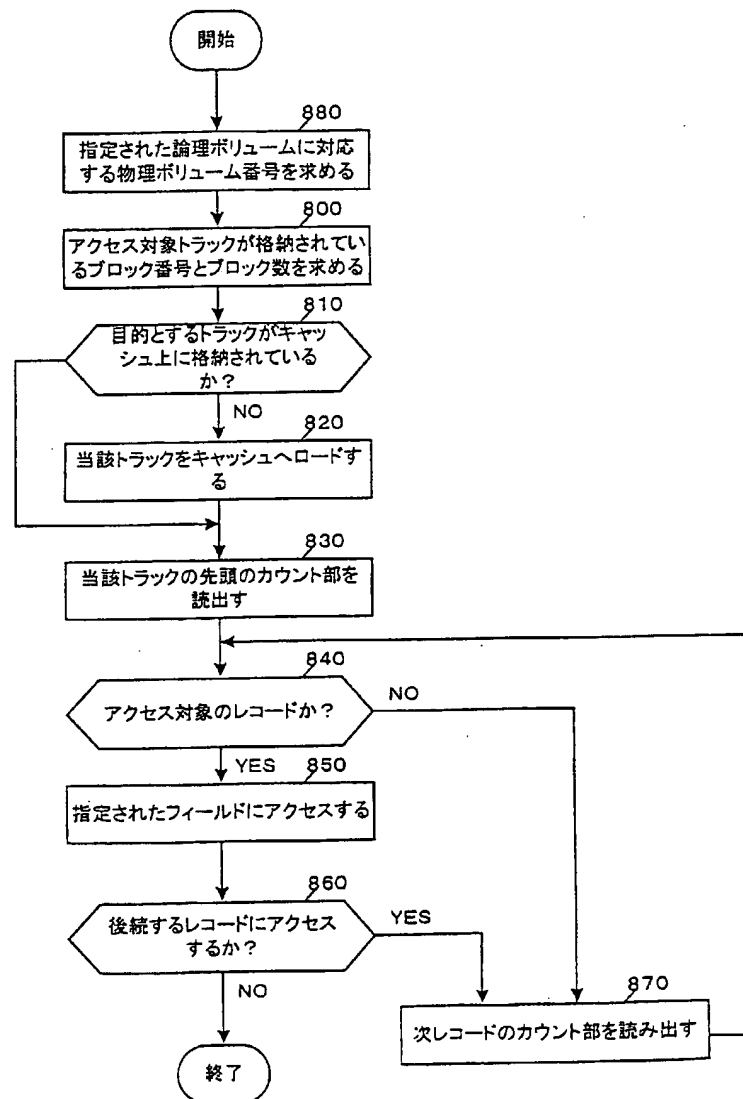


第 6 図



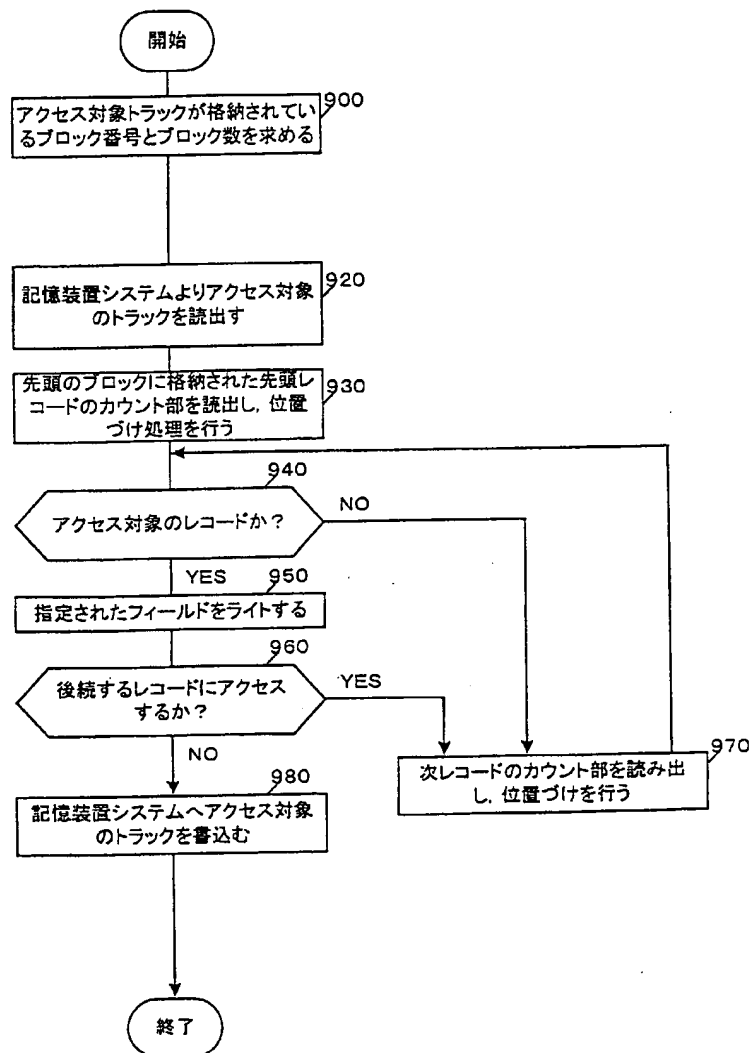
7 / 9

第 7 図



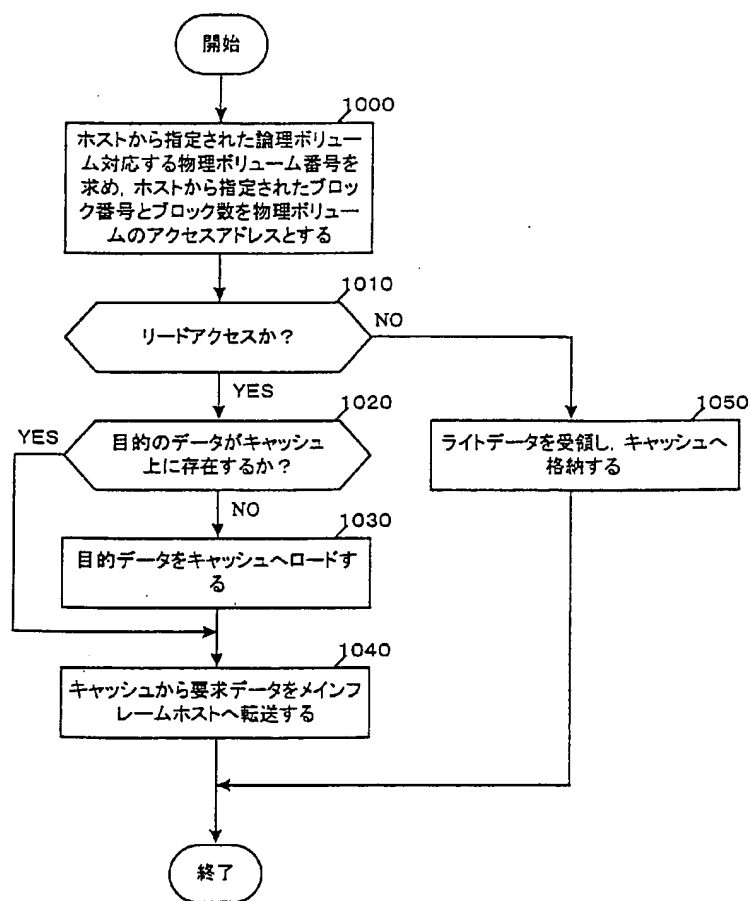
8 / 9

第 8 図



9 / 9

第 9 図



## INTERNATIONAL SEARCH REPORT

 International application No.  
 PCT/JP98/03335

 A. CLASSIFICATION OF SUBJECT MATTER  
 Int.Cl.<sup>6</sup> G06F12/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

 Minimum documentation searched (classification system followed by classification symbols)  
 Int.Cl.<sup>6</sup> G06F12/00, G06F3/06, G11B20/10

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP, 6-12786, A (Hitachi, Ltd.), 21 January, 1994 (21. 01. 94),	1-4
Y	JP, 7-168675, A (International Business Machines Corp.), 4 July, 1995 (04. 07. 95), & US, 5459853, A	1-4
Y	JP, 8-241174, A (Fujitsu Ltd.), 17 September, 1996 (17. 09. 96)	1-4
Y	JP, 9-258908, A (Hitachi, Ltd.), 3 October, 1997 (03. 10. 97), & EP, 785500, A1	1-4
Y	JP, 10-11223, A (NEC Ibaraki Ltd.), 16 January, 1998 (16. 01. 98)	1-4
Y	JP, 10-112136, A (NEC Ibaraki Ltd.), 28 April, 1998 (28. 04. 98)	1-4

☒ Further documents are listed in the continuation of Box C.
 ☐ See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

 Date of the actual completion of the international search  
 14 October, 1998 (14. 10. 98)

 Date of mailing of the international search report  
 27 October, 1998 (27. 10. 98)

 Name and mailing address of the ISA/  
 Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP98/03335

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Norimitsu Furukatsu, et al., "Background of Developoment of Parallel ACOS Series and its Outline (in Japanese)", p.4-10, NEC Technical Journal, <u>48</u> [9] (25-09-1995)	1-4

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	J P, 10-112136, A (茨城日本電気株式会社), 28. 4月. 1998 (28. 04. 1998)	1-4
Y	NEC技法、48 [9] (25-09-1995) 古勝紀誠 他「パラレルAC OSシリーズの開発背景と概説」p.4-10	1-4

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☒ **FADED TEXT OR DRAWING**
- ☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**